

기획서

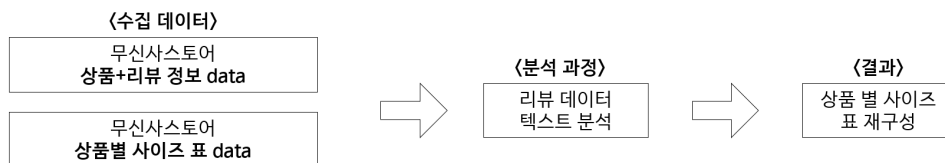
1) 주제

- 구매자의 선호 사이즈에 맞는 무신사 스토어 리뷰 기반 의류 추천 시스템

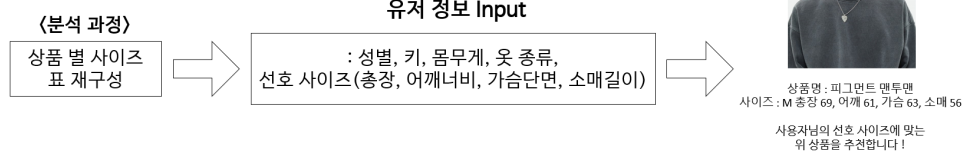
2) 분석 목적

- 사이즈 표가 사이트에 명시 되어 있음에도 불구하고 사이즈 미스 반품 사례가 여전히 발생하고 있음
- 옷 마다 실제 사이즈 다르니 리뷰 분석해서 원래 게시되어 있는 리뷰 기반한 실제 사이즈 예측하고 이를 기반으로 상품을 추천하고자 함

1. 시스템 구축 과정



2. 시스템 실현 과정



3) 선정된 주제 이슈 조사

- "쇼핑몰 옷 사이즈 애매해 불편, 개선해달라" 청원

"쇼핑몰 옷 사이즈 애매해 불편, 개선해달라" 청원

표준의류사이즈표에도 불구하고 제각각인 쇼핑몰별 옷 사이즈로 인해 소비자들이 불편을 호소하고 있다. 이에 국내 의류 표준 규격 사이즈를 정해달라는 청원이 등장했다. 3일 청와대 국민청원 게시판에 '국내 의류 표준 규격 사이즈 관련 법을 제정해주십시오'라는 제목의

<https://www.ekoreanews.co.kr/news/articleView.html?idxno=42138>

— 청원진행중 —

국내 의류 표준 규격 사이즈 관련 법을 제정해
주십시오

참여인원 : [3,507명]

카테고리 기타 청원시작 2020-02-03 청원마감 2020-03-04 청원인 never***

- 패션 상품의 반품의 비중은 전체 반품 상품 중 가장 높게 나타났는데 소비자의 45%가 사이즈, 핏, 컬러의 문제로 반품

www.sfti.or.kr

https://www.sfti.or.kr/pdf_files/pdf_file/25_3_02_280_290_23_663.pdf

- '들쭉날쭉' 옷 사이즈... '골치 아픈' 소비자

'들쭉날쭉' 옷 사이즈... '골치 아픈' 소비자

#. A씨는 최근 온라인 쇼핑몰에서 구매한 셔츠를 환불했다. M사이즈를 샀는데 다른 브랜드보다 사이즈가 작게 나온 것이다. 쇼핑몰에 항의할까 생각도 했지만 이내 마음을 접었다. 사이즈로 인한 교환 요청은 '단순 번심'처럼 고객의 책임으로 처리된다는 것을 알고 있기 때

<https://www.fnnews.com/news/202002251738429663>



- ⇒ 쇼핑몰의 옷 사이즈에 대한 모호한 기준이 의류를 구매하는데 있어, 불편함을 제공함

4) 데이터 준비

- 의류 데이터 (**출처: 무신사 크롤링**)
- 유저 정보
 - 유저ID / 성별 / 키 / 몸무게 / 옷 종류(전체, 세분화) / 사이즈(선호하는 어깨너비, 가슴단면, 소매길이, 총장) / (색상)
 - 키와 몸무게가 설정되어있지 않은 리뷰는 제외
 - 오버핏같은 경우 본인의 사이즈가 아닌 선호하는 사이즈를 넣음으로 해결 가능
- 색상
 - 크롤링을 통해 수집한 데이터의 상품 이름에서 '색상' 추출
 - 방법1) input값에 색상 O : 최종적으로 군집화 된 데이터 내에서 의상 추천
 - 방법2) input값에 색상 X : 최종적으로 의상 추천할 때 기준으로 가중치 부여
- 순위
 - 사이트 내 판매순 1년 정렬하여 상품 정보 추출
- 옷 종류
 - 남성 상의로 한정
 - 카테고리 별로 의류 정보 수집
 - input으로 옷 카테고리 입력(전체, 신상제외 항목 5개)
 - 전체 카테고리를 넣으면 어떤 카테고리의 옷을 살지 고민하는 사람에게도 도움이 될 것이라 예상
- 상품 별 사이즈 표
 - 사이트에서 상품별로 제공하는 사이즈 표의 값
- 리뷰 텍스트
 - 카테고리 5개 (긴소매, 반소매, 맨투맨/스웨트셔츠, 후드티, 셔츠/블라우스) 100개 상품에 대해서 100개씩 크롤링하여 5만개 수집 예정
 - 상품사진 리뷰 50개 + 일반리뷰 50개
 - 리뷰 유용한순, 사이즈 필터 선택하여 사이즈 관련한 리뷰를 가져오기 위해 필터하여 수집
 - "도움돼요" 개수
 - 사이즈 계산 시 가중치 부여 요소로 활용 가능
 - "도움돼요" 개수가 0개인 상품들은 순서가 무작위로 배치되는 가능성 존재
 - 상품 사이즈의 리뷰가 적혀있기 때문에 우선 상위 100개 수집



- 수집 기준일 명시
 - 날짜가 지날수록 순서가 변동될 수 있음
- 약 5만개의 데이터를 통해 워크플로우 구축 후 데이터 추가 수집 예정

5) 분석 과정

1. 사이즈 표 재구성

- 무신사 스토어에서 상품별로 사이즈 표 제공함 → 제품별 사이즈 표를 리뷰를 기반으로 재구성 → 사이즈 재구성 데이터(모델) 구축

2. 사용자의 선호도 사이즈 input을 받고, 1번 과정을 통해 재구성된 표 → 상품, 사이즈 추천


- input
 - 성별, 키, 몸무게, 옷 종류, 선호 사이즈(총장, 어깨너비, 가슴단면, 소매길이)
- output
 - 상품명, 추천 사이즈, 추천 멘트

⇒ 결론 : 1번 (사이즈 표 재구성) → 2번 사용자 input → 상품, 사이즈 추천

6) 분석 사례 조사

- 리뷰 기반 의류 사이즈 추천 시스템 프로젝트

제 17회 보아즈(BOAZ) 빅데이터 컨퍼런스 - [SiZoAH] : 리뷰 기반 의류 사이즈 추천시스템
 제 17회 보아즈(BOAZ) 빅데이터 컨퍼런스 - [SiZoAH] : 리뷰 기반 의류 사이즈 추천시스템 -
 Download as a PDF or view online for free

 <https://www.slideshare.net/slideshow/17-boaz-sizoah/255973431>

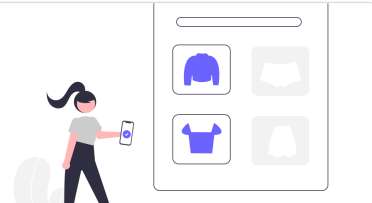


- 소비자의 사이즈 고민을 해결하기 위한 가장 완벽한 방법

소비자의 사이즈 고민을 해결하기 위한 가장 완벽한 방법

크리마 핏 | 어느 때보다 온라인 쇼핑이 활성화되었음에도 불구하고, 소비자들은 여전히 본인에게 맞는 사이즈, 원하는 핏감의 사이즈 선택을 어려워합니다. 왜 그럴까요? 보통 소비자는 온라인에서 의류 상품을 구매할 때, 구매 리뷰와 상품 실측 치수를 통해 나와 비슷한

 <https://brunch.co.kr/@cremasolution/24>



7) 주차별 일정(6.3~6.27)

- 1주차
 - 기획서 선정 및 작성, 데이터 수집 및 EDA, 데이터set 구축
- 2주차
 - 텍스트 전처리, 사이즈 군집화
- 3주차
 - 모델링
- 4주차
 - ppt 작성 및 정리, 보고서 작성

Aa 이름	📅 날짜	☰ 태그
<u>주제 선정/기획서 작성</u>	@2024년 5월 23일 → 2024년 6월 4일	
<u>데이터 수집 및 EDA</u>	@2024년 6월 5일 → 2024년 6월 7일	
<u>피드백</u>	@2024년 6월 5일	
<u>효감</u>	@2024년 6월 6일	
<u>모델링</u>	@2024년 6월 16일 → 2024년 6월 23일	
<u>ppt 작성 및 정리, 보고서 작성</u>	@2024년 6월 24일 → 2024년 6월 26일	
<u>발표</u>	@2024년 6월 27일	
<u>텍스트 전처리 및 사이즈 군집화</u>	@2024년 6월 8일 → 2024년 6월 15일	
<u>중간점검</u>	@2024년 6월 17일	
<u>기획서 제출</u>	@2024년 6월 7일	

8) 팀원 소개

- 류수연
 - 데이터 수집 / 데이터 전처리 / 모델링 / 결과 분석

- 모델링, EDA 및 시각화, 결과 분석, 발표자료 제작
- 성지현
 - 데이터 수집 / 데이터 전처리 / 모델링 / 결과 분석
 - 데이터 수집, EDA 및 시각화, 결과 분석
- 심태양
 - 데이터 수집 / 데이터 전처리 / 모델링 / 결과 분석
 - 작업 플로우 구축, 데이터 전처리, 결과 분석

9) 사용 기술

- 언어
 - python
- 작업 환경
 - Google Colab
- 패키지
 - numpy / pandas / sklearn / pytorch
- 기록 및 회의
 - Notion, Discord